



Journal of Educational Sciences

Journal homepage: <https://jes.ejournal.unri.ac.id/index.php/JES>



P-ISSN
2581-1657

E-ISSN
2581-2203

Leveraging Artificial Intelligence (AI) to Enhance Teachers' Spoken English at an Islamic Boarding School

Ainol Mardhiah*¹, Ibnu Hajar¹, Diana Purwati², Hayatul Muna^{3,4}, Riska Ariani¹

¹Department of English Education, Universitas Bumi Persada, Lhokseumawe, 24351, Indonesia

²Department of English Education, STKIP Yapis Dompur, Nusa Tenggara Barat, 84251, Indonesia

³Department of English Education, UIN Sultanah Nahrasyiah, Lhokseumawe, 24351, Indonesia

⁴Eka Learning Center (ELC), Lhokseumawe, 24351, Indonesia

ARTICLE INFO

Article history:

Received: 19 Sept 2025

Revised: 13 Oct 2025

Accepted: 25 Oct 2025

Published online: 05 Nov 2025

Keywords:

Artificial Intelligence (ai),
Boarding School,
English,
Speechace,
Speaking for Daily Communication

* Corresponding author:

E-mail: ainolmardhiah@unbp.ac.id

Article Doi:

<https://doi.org/10.31258/jes.9.6.p.5163-5176>

This is an open access article under the [CC BY-NC-SA](https://creativecommons.org/licenses/by-nc-sa/4.0/) license.



ABSTRACT

Many Indonesian *pesantren* (Islamic boarding schools) adopt bilingual policies requiring daily English communication, yet limited teacher proficiency often hinders their effective implementation. This study investigates the use of Artificial Intelligence (AI) SpeechAce to enhance teachers' speaking skills for daily communication in a *pesantren* in Lhokseumawe, Aceh, Indonesia. Using a mixed-method explanatory sequential design, 29 teachers participated in a one-month intervention consisting of AI-based practice and training, pre- and post-tests, reflective logbook, and semi-structured interviews. Quantitative results revealed consistent improvement across pronunciation, fluency, vocabulary, and grammar, with vocabulary and grammar showing the greatest gains. Qualitative findings highlighted teachers' positive perceptions of AI in promoting self-regulated learning and as an anxiety-reducing tool for speaking practice, despite challenges such as technical issues and the need for a longer and a more structured guidance from a human-teacher. While AI SpeechAce effectively enhanced teachers' speaking confidence through safe, feedback-driven practice relevant to their daily communication, its limitations shaped by linguistic gaps and technical challenges, underscore the need for future blended models and extended interventions to ensure stronger transfer into real-world communicative spoken English skills.

1. Introduction

Many *pesantren*, Islamic boarding schools, in Indonesia, have adopted bilingual policies which require students to use English in daily communication both inside and outside the classroom. Thus, the ability to communicate in English, particularly in speaking, is an essential skill for students in *pesantren* that aspire to produce graduates that can compete globally (Rahman et al., 2023). However, in achieving this goal, *pesantren* often encounters challenges due to the limited speaking proficiency of teachers (Amiruddin & Jannah, 2019; Masrukhan et al., 2023), who

ideally should serve as role models and sources of meaningful input in English. In the case of an Islamic Boarding school located in Lhokseumawe, preliminary findings indicate that the majority of teachers possess only basic or upper beginner speaking abilities.

Such a finding has a significant impact on the language input and models provided to students, which ultimately affect their motivation to use English in daily interactions at the *pesantren*. Teachers also reported that they rarely observe students speaking English in class, let alone outside the classroom. This creates a gap between the institution's vision and its practical implementation. In response to this challenge, the leaders and stakeholders of the *pesantren* have taken the initiative to conduct language training for teachers. However, due to time constraints and workloads, the training is typically held only once or twice per year. In fact, consistent and continuous practice is essential for improving speaking proficiency (Alam et al., 2024). If this issue is not addressed, the *pesantren*'s vision of producing graduates who are proficient in English will be difficult to achieve. Hence, there is an urgent need to enhance teachers' speaking abilities in continuous, flexible, and effective ways.

Recent studies have revealed the potential of Artificial Intelligence (AI) to enhance speaking skills in second or foreign language contexts. Mardhiah et al., (2024) highlight the benefits of AI tools in improving pronunciation, fluency, vocabulary, and grammar accuracy across various contexts. Similarly, AI tools integrated with social network interactions has also been found to increase learners' speaking practice and classroom participation (Zou et al., 2023). A study in Egypt further confirm that AI-based practice could improve linguistic output, although its role in reducing foreign language anxiety remains limited (El Shazly, 2021). Collectively, these findings indicate that AI can serve as an effective tool for enhancing speaking skills by providing constructive feedback and continuous practice opportunities.

AI Speech recognition and pronunciation training tools, such as ELSA Speak and SpeechAce are highly valuable in contexts where students have limited opportunities for regular interaction with native speakers or engaging in speaking practice. These platforms provide detailed feedback which enable users to reflect on pronunciation, rhythm, and intonation (Karim et al., 2023; Zhang et al., 2023). A study on the use of Speechace in an Indonesian higher education also showed improvements in students' speaking performance, particularly among female students in everyday communication tasks (Ningsih, 2024). SpeechAce was then selected over other tools because it has been empirically tested in the Indonesian context and provides more structured and real-time feedback that enable users to monitor their speaking skills and reduce anxiety (Yeh, 2024; Dong et al., 2024). Within the scope of this study, the use of SpeechAce provides teachers with flexible opportunities to practice anytime and anywhere, serving as a practical intervention to enhance their speaking for daily communication.

Despite the promising results of AI utilization in enhancing speaking proficiency, current studies remain limited in scope as many existing research focuses on students and at university level. Meanwhile, teachers, who play a central role in

sustaining bilingual policies, have received far less attention. Moreover, although AI tools have demonstrated their effectiveness in enhancing speaking abilities, their integration into the specific context of daily communication in *pesantren* settings has not yet been investigated. These gaps highlight the need to explore AI-based practice into teacher development, especially in the contexts where English serves not only as an academic subject but also as a medium of daily interaction.

This study aims to address these gaps by leveraging AI to enhance speaking for daily communication of teachers at an Islamic boarding school in Lhokseumawe. The research focuses on two main questions: (1) Is AI SpeechAce effective in enhancing teachers' speaking ability for daily communication at a *pesantren* in Lhokseumawe? and (2) How effective is AI SpeechAce in improving teachers' speaking for daily communication at a *pesantren* in Lhokseumawe?. Using a mixed-methods approach, this study aims not only to measure improvements through pre-tests and post-tests but also to investigate teachers' experiences through interviews and reflective log books. The novelty of this research lies in its focus on teachers as the primary agents of bilingual policy at the *pesantren*, the integration of AI into *pesantren* contexts, and the development of AI-based practices that can be used individually, flexibly, and continuously by teachers.

2. Methodology

This study was conducted at an Islamic boarding school located in Lhokseumawe, Aceh, Indonesia. A mixed-method approach with an explanatory sequential design, as developed by Creswell (1999), was employed. This design was chosen because it provides a comprehensive understanding of the effectiveness of AI SpeechAce by first collecting and analyzing quantitative data, followed by qualitative data to explain and elaborate on the findings. The participants were selected using purposive sampling. Not all teachers at the *pesantren* were included in the study. Only teachers directly involved in supporting students' English use such as language teachers and dormitory teachers were selected. This selection was based on their intensive role in accompanying students both in classrooms and in dormitory life. In total, 29 teachers participated in the study, consisting of 13 language teachers and 16 dormitory teachers.

Both quantitative and qualitative data were collected. The quantitative phase aimed to measure the effectiveness of AI in improving teachers' speaking ability. Pre-tests and post-tests of each of the units were conducted using AI SpeechAce, which automatically assessed participants' speaking performance and generated real-time scores and feedbacks across four aspects: pronunciation, fluency, vocabulary, and grammar. The intervention lasted one month, consisting of face-to-face pre-tests, structured training sessions, and post-tests. Each session ran for approximately 90 minutes, held twice per week. The training was organized to combine direct instruction, AI-based practice, and reflection on feedback. The detailed schedule of this research intervention can be seen below:

Table 1. Detailed Schedule of AI SpeechAce Intervention

Day and Date	Activities
Saturday, 09/08/ 2025	Intervension 1 - Pre-Tests & Log Book Distribution
Sunday, 10/08/ 2025	Intervension 2 - Self - Introduction
Saturday, 16/ 08/ 2025	Intervension 3 - Daily activity
Sunday, 17/ 08/2025	Intervension 4 - Possession
Saturday, 23/ 08/ 2025	Intervension 5 - Asking Permission & Giving Suggestion
Sunday, 24/ 08/ 2025	Intervension 6 - Describing food
Saturday, 30/ 08/ 2025	Intervension 7 - Describing Food & Log Book Submission
Sunday, 31/ 08/ 2025	Intervension 8 - Post Tests

In addition to these sessions, teachers were required to complete independent AI practice using the SpeechAce platform. This independent work ensured consistent exposure and active engagement with the technology. Throughout the intervention, teachers systematically documented their experiences, challenges, and progress in reflective logbooks. They also shared regular updates via WhatsApp which allow researchers to close monitoring. To enrich the quantitative findings, semi-structured interviews were conducted after the intervention. These interviews provided valuable insights into teachers' experiences with the AI SpeechAce.

Quantitative data from the pre-tests and post-tests, which were scored by AI Speechace using the IELTS (International English Language Testing System) 0-9 band scale, were analyzed using descriptive statistics. Average scores and score distributions were calculated to identify trends in speaking improvement. Each aspect of speaking such as pronunciation, fluency, vocabulary, and grammar, was also analyzed separately to examine specific areas of progress. The results were then presented in a table form to illustrate learning trends clearly. For qualitative data, information from interviews and log books was reduced, categorized, and analyzed thematically to identify major themes regarding the role of AI and the challenges teachers faced.

To ensure the validity and reliability of the findings, source triangulation was systematically conducted by comparing results from the pre-tests and post-tests, interviews, and reflective logbooks. This triangulation allowed the data to be cross-verified from multiple perspectives. Member checking was also applied to allow teachers to review and confirm the accuracy of the interpretations of their interview and log book data. To ensure consistency and strengthen the credibility of this research conclusions, both quantitative and qualitative findings were carefully analysed.

3. Results and Discussion

AI SpeechAce in enhancing speaking for daily communication

Pre-test and post-test scores were collected across four instructional training units: *Self Introduction and Daily Activities, Possession, Asking Permission and Giving Suggestion, and Describing Food*. Each unit was assessed based on four key aspects

of speaking skills: pronunciation, fluency, vocabulary, and grammar. Scores from both assessments were systematically compared to capture measurable progress. The results are presented in tables to clearly display the changes in average scores across units. Descriptive analyses were conducted to further explain patterns of improvement and areas of challenge. This approach provides a comprehensive understanding of the impact of SpeechAce on enhancing the teachers’ spoken English skills. These findings directly address Research Question 1 of this study.

The analysis of Unit 1 shows a clear improvement in teachers’ speaking daily communication skills related to the topic of Self-introduction and daily activities. The average overall score increased from 4.97 to 5.48. Vocabulary showed the greatest increase (from 4.9 to 5.6), followed by grammar (from 4.8 to 5.4). Pronunciation improved from 5.0 to 5.5, while fluency increased slightly from 5.2 to 5.4. These results (shown in the table below) suggest that the intervention was particularly beneficial in developing teachers’ vocabulary and grammatical accuracy, with moderate improvements in pronunciation and fluency.

Table 2. Test Results of Unit 1 (Self Introduction and Daily Activity)

Participant	P	F	V	G	Pre-Test Score	P	F	V	G	Post-Test Score
AR	1,5	1,5	1,5	1,5	1,5	6	6,5	5,5	5,5	6
A	3,5	3,5	3	3,5	3,5	5	5	5	5	5
HM	1,5	1,5	1,5	1,5	1,5	5,5	5,5	6	5,1	4
HR	5,5	5,5	6	5,5	5,5	5,5	5,5	5,5	6	5,5
IN	6,5	6,5	5,5	5,5	6	6	6	5,5	5,5	6
KA	6	6	5	5	4,5	6,5	6,5	5	5	5
LE	6	6	6	5,5	6	5,5	5,5	5	6	5,5
MR	1,5	1,5	1,5	1,5	1,5	5,5	5,5	6,5	6	6
M	6	6	5,5	0,5	6	3,5	3,5	3	3,5	3
MZ	3	3,5	3,5	3,5	3	4,5	5	4	5	4
NL	5	5	5	5,5	5	5	5	6	5,5	5,5
NTS	6,5	6,5	6	5,5	6	6	6	5,5	6	6
NSN	5,5	5,5	5,5	6	5,5	6,5	6,5	6,5	6,5	6,5
NF	2	2	1,5	1,5	1,5	6,5	6,5	5	5,5	5,5
NA	6,5	6,5	5,5	5	5	6,5	6	5,5	5,5	4,5
NO	5,5	5,5	6	5,5	6,5	6,5	6,5	6,5	6	6,5
PR	5,5	5,5	4,5	5,5	5	5,5	5,5	5	5	5
RF	5,5	5,5	4,5	5	5	6,5	6,5	6	6,5	6,5
SI	3	3	3	3	3	3,5	3,5	3	3,5	3
SA	6	6	5,5	5,5	5,5	6	6	5,5	6	6
S	6,5	6	6	6,5	6,5	6,5	6,5	6	6	6,5
UZ	3	3,5	3	3,5	3	3	3,5	3,5	3,5	3,5
UF	7	6,5	6	5,5	6,5	7	6,5	5,5	6	6
WI	3	3,5	3,5	3,5	3,5	5	5	4,5	5	4,5
WA	5	5	4	5	5	6	6	4,5	5,5	5,5
FFA	5,5	5,5	5	5,5	5	5	5,5	5,5	5,5	5,5
HG	3	3,5	3	3,5	3,5	5,5	5,5	4,5	5	4,5
MN	2	2	1,5	1,5	1,5	3	3,5	3	3,5	3,5
SR	6	6	5	5	5	6,5	6	5,5	6	6
Average	4,5	4,5	4,2	4,1	4,3	5,5	5,5	5,1	5,3	5,1

The results of Unit 2, which can be seen in the table below, reveal a positive trend in teachers' performance from pre-test to post-test. For the topic related to possession, the average score increased from 4.67 to 5.34, indicating measurable improvement. Vocabulary showed the highest rise (from 4.6 to 5.5), followed by grammar (from 4.9 to 5.6). Pronunciation (from 5.2 to 5.6) and fluency (from 5.4 to 5.7) also improved, though at a more modest rate. These findings suggest that the intervention was especially effective in enhancing vocabulary and grammar while supporting gradual progress in pronunciation and fluency.

Table 3. Test Results of Unit 2 (Possession)

Participant	P	F	V	G	Pre-Test Score	P	F	V	G	Post-Test Score
AR	6,5	6	4,5	5	4,5	6	6	5	5,5	5,5
A	5	5	4,5	5	5	5	5	5	5,5	5
HM	5	6,5	6	5	4,5	5	5,5	5	5	4,5
HR	5	5	4	4,5	4	4,5	5,5	6	6	5,5
IN	3,5	5,5	5,5	6	5	7	6,5	6	7	6,5
KA	6	6	5,5	5	4	6	6	5,5	5	4
LE	5,5	5,5	5,5	6	5,5	5	5	5	6	5,5
MR	5	5	4	5,5	5	5,5	5,5	6	6	5,5
M	5	5	4,5	5	5	5	5	5,5	5	5
MZ	5	5,5	6	5,5	5,5	5	5,5	5	5,5	4
NL	1,5	1,5	1,5	2	1,5	5	5	4,5	5,5	5
NTS	6,5	6	6	6,5	6	7	6,5	6	6	6,5
NSN	5,5	5,5	5,5	6	5,5	6	6	6,5	6,5	6
NF	6	6	5,5	5	5	5,5	5,5	5,5	5	5
NA	6,5	6,5	4	4,5	4	5,5	6	7,5	5	4
NO	7	6,5	6,5	6	6,5	7	7	6,5	6	7
PR	5	5	4,5	5,5	5	5	5,5	4,5	5,5	5
RF	6,5	6,5	4,5	5	5,5	7	7	6	6	6,5
SI	4,5	5	5	5	4	3,5	3,5	3,5	3	2,5
SA	2,5	3,5	3	3,5	3	5,5	5,5	5,5	5	5,5
S	7	7	5,5	6	6,5	7	6,5	6,5	7	7
UZ	3,5	4	3	3	2,5	5	5	4,5	5	4,5
UF	6,5	6	5,5	6	6	7	6,5	6,5	7	7
WI	5	5	4	5	5	5	5	5	5	5
WA	7	7	5,5	5	6	6,5	6,5	6	6	6
FFA	5,5	5,5	4,5	5	4,5	5	5,5	5,5	5,5	5,5
HG	5,5	5,5	4,5	5	4,5	5,5	5,5	4,5	5,5	5
MN	2	2	1,5	1,5	1,5	4,5	4,5	4,5	5	5,5
SR	7	6,5	4,5	5	5	6	6	5,5	6	5,5
Average	5,2	5,4	4,6	4,9	4,6	5,6	5,7	5,5	5,6	5,3

Unit 3 results demonstrate teachers' steady progress in speaking related to the topic of Asking Permission & Giving Suggestion, with the average score rising from 5.17 to 5.53. Vocabulary recorded the highest improvement (from 5.1 to 5.6), followed by grammar (from 5.0 to 5.5). Pronunciation increased from 5.2 to 5.5, and fluency from 5.4 to 5.5. Although the gains were relatively modest compared to earlier units, the consistent increase across all aspects indicates that participants continued to benefit from AI SpeechAce. These findings indicate that the tool plays role as a supportive learning aid, particularly in enhancing teachers' speaking skills.

However, its contribution in this stage appears incremental rather than transformative. The results can be seen as follows:

Table 4. Test Results of Unit 3 (Asking Permission & Giving Suggestion)

Participant	P	F	V	G	Pre-Test Score	P	F	V	G	Post-Test Score
AR	6,5	6	4,5	5	4	7	6,5	4,5	5	4
A	5	5	4,5	4,5	4,5	5	5	5	5,5	5
HM	4,5	5,5	5	4,5	4	5,5	5,5	4,5	4,5	4
HR	4,5	5,5	6,5	7	6	4,5	5,5	6,5	7	6
IN	6,5	6,5	6,5	7	6,5	6,5	6,5	6,5	7	6,5
KA	6	6	5	4,5	4	6	6	4,5	4,5	4
LE	5	5	5	6	5,5	5	5	5	6	5,5
MR	5	5	5	5,5	5	5	5	5	5,5	5
M	5,5	5,5	5,5	4,5	5	5,5	5,5	5,5	4,5	5
MZ	4,5	5,5	6,5	5	5,5	4,5	5	5	5	4,5
NL	5,5	5,5	5	5,5	5,5	5,5	5,5	5	5,5	5,5
NTS	6,5	6	5,5	6	6	6,5	6	6,5	6,5	6,5
NSN	5,5	5,5	5,5	6	5,5	5,5	5,5	5,5	6	5,5
NF	4,5	4,5	3	3,5	2,5	4,5	4,5	3	3,5	2,5
NA	7	6,5	5	5	4	7	6,5	5	5	4
NO	3,5	3,5	4,5	5	4	3,5	3,5	4,5	5	4
PR	5,5	5,5	5	6	5,5	5,5	5,5	5,5	6	5,5
RF	7	7	5,5	6	6,5	7	6,5	7	7	7
SI	3	3,5	3,5	3	3,5	5,5	5,5	5	5	4
SA	5	5,5	4,5	5	5	5	5,5	4,5	5	5
S	7,5	6,5	5,5	6	6,5	7,5	6,5	5,5	6	6,5
UZ	3	3,5	4,5	3,5	3	5,5	5,5	5	5,5	5,5
UF	7	6,5	5,5	5,5	6	7	6,5	5,5	5,5	6
WI	3	3,5	3	3,5	3,5	3	3,5	3	3,5	3,5
WA	7,5	7	5	5,5	6,5	7,5	7	5	5,5	6,5
FFA	5,5	5,5	5	5	5	5	5,5	5,5	5,5	5,5
HG	2	2	1,5	2	2	5,5	5,5	4,5	5	4,5
MN	3	3,5	3,5	3	3,5	5	5	4,5	5,5	5
SR	7,5	7	4,5	6	6,5	7,5	7	4,5	6	6,5
Average	5,2	5,3	4,8	5	4,8	5,64	5,6	5,1	5,4	5,1

The analysis of Unit 4 also shows improvement in teachers’ speaking for daily communication related to the topic of Describing Food, with the overall average rising from 5.43 to 5.70. Vocabulary demonstrated the most noticeable increase (from 5.7 to 6.0), followed by grammar (from 5.6 to 5.8). Pronunciation (5.7 to 5.8) and fluency (5.8 to 5.8) remained relatively stable, suggesting that teachers showed their progress in these areas. The results demonstrated in the table below highlight the role of AI SpeechAce in enhancing vocabulary and grammatical control when describing food, while pronunciation and fluency remained stagnant.

Table 5. Test Results of Unit 4 (Describing Food)

Participant	P	F	V	G	Pre-Test Score	P	F	V	G	Post-Test Score
AR	5,5	5,5	6	6	5,5	5,5	5,5	6	6	5,5
A	4	5	6	6	5	4,5	4,5	5	5	5

HM	4,5	5,5	5,5	5	5	5	6	7	5	5,5
HR	5	5,5	5,5	5,5	4,5	4,5	5,5	5,5	6	5,5
IN	6	6	7,5	6,5	6,5	6	6	7,5	7	6,5
KA	6	6	6,5	5,5	5	5	5,5	6	5	5
LE	6	6	5,5	6	6	6	6	5,5	6	6
MR	6	6	5,5	6	6	6	6	5,5	6	6
M	6,5	6	5	5	5	5,5	5,5	5	5	5
MZ	5	5,5	6,5	5,5	5	5	5,5	6,5	5,5	5
NL	5,5	6	5,5	6	6	6	6	6,5	5,5	6
NTS	5,5	6	5,5	6	6	7	6,5	7	7,5	7
NSN	5,5	5,5	5	6,5	5,5	5,5	5,5	6,5	6,6	6
NF	6	6	5	5	5	5,5	5,5	6	5	5,5
NA	5,5	5,5	6	4,5	4,5	5,5	6	6,5	5	5,5
NO	7	6,5	5,5	6	6	7	6	5,5	5,5	6
PR	6,5	6,5	5,5	6	6	6	6	5,5	6	6
RF	6,5	6,5	6	6,5	6,5	6,5	6,5	7	7	6,5
SI	5	5,5	6,5	5	4,5	4,5	5	6,5	5	4,5
SA	6,5	6,5	6,5	6,5	6,5	6	6	6,5	6,5	6,5
S	6	6	5,5	5	5	6	6	6,5	6,5	6
UZ	4,5	5	5,5	5,5	5	5,5	5,5	6	5	5,5
UF	7	6,5	7	6,5	6,5	7	6,5	7	6,5	6,5
WI	4	5	5	4	4,5	5	5	5	5	4,5
WA	6,5	6	5	6	6	6,5	6	5,5	6	6
FFA	5,5	5,5	5,5	5	5	6	6	6	6	6
HG	6	6	5	5	5,5	6	6	5	5	5,5
MN	4	4,5	5	5	4	4	4,5	5	5	4
SR	6,5	6	5,5	6	6	6,5	6,5	6	6	6,5
Average	5,7	5,8	5,7	5,6	5,4	5,7	5,8	6,0	5,8	5,6

Based on the results across the four instructional units, it can be clearly seen that there is an improvement in teachers' speaking skills, though the degree of improvement varied depending on the unit. The largest overall was in Unit 2 (Possession), where the average score rose by 0.67 points. Unit 1 (Self Introduction & Daily Activities) also showed a notable gain of 0.59 points, particularly because participants started at a lower baseline in this area. By contrast, Unit 3 (Asking Permission & Giving Suggestion) and Unit 4 (Describing Food) demonstrated more modest increases of 0.28 and 0.27 points, respectively. These smaller improvements suggest that as tasks grew more complex and context-specific, teachers' progress slowed, highlighting the need for additional instruction when dealing with advanced communicative functions.

When comparing across speaking aspects, vocabulary consistently showed the most significant gains. In every unit, teachers expanded their vocabulary, with improvements averaging around 0.35–0.40 points. Grammar followed closely, showing steady growth of about 0.30–0.40 points, especially in Units 1 and 2 where sentence structures were relatively straightforward. Fluency, however, improved at a more moderate pace, with gains ranging from 0.1 to 0.3 points. This indicates that while teachers gained confidence, they still encountered difficulties in maintaining smooth delivery. Pronunciation revealed the most uneven pattern. Noticeable progress was observed in Units 1 and 2, where language was more predictable, but scores plateaued in Units 3 and 4. This suggests that the AI SpeechAce corrective

feedback alone was not sufficient to address pronunciation challenges. This trend is summarized in Table 6:

Table 6. Comparison of Pre-Test and Post-Test Results (Units 1–4)

Unit	Topic	Pre-Test Avg.	Post-Test Avg.	Improvement	Main Trend
1	Self Introduction & Daily Activity	4.52	5.11	+0.59	Strong improvement, especially in Fluency and Grammar.
2	Possession	4.67	5.34	+0.67	Highest gain; good progress in Pronunciation and Fluency.
3	Permission & Suggestion	4.84	5.12	+0.28	Slight improvement; challenges with modal expressions.
4	Describing Food	5.43	5.70	+0.27	Modest progress; growth mainly in Vocabulary.

Taken together, these findings highlight the specific strengths and limitations of using AI SpeechAce in enhancing speaking for daily communication skills of teachers at an Islamic Boarding School in Aceh. On the one hand, the intervention was highly effective in enriching teachers' vocabulary and grammar, providing them with the resources and confidence for daily communicative functions. On the other hand, the relatively smaller gains in fluency and pronunciation suggest that AI SpeechAce feedback needs to be complemented with more interactive, human-centered speaking practice. Moreover, the stagnant results observed in later units underscores the importance of carefully scaffolding learning tasks, moving teachers gradually from structured practice toward more complex, real-life communicative situations. To investigate further, the qualitative data is presented below.

Factors Contributing to the Success of AI Implementation

The data reveal several key factors that contributed to the positive impact of the SpeechAce intervention. A predominant theme across the responses is a significant increase in self-confidence. Participants frequently reported feeling more courageous and less embarrassed when speaking English. The AI environment was consistently described as a "safe space" free from the judgment of human listeners (e.g., Respondent RF: "*bisa berbicara sendiri dengan AI tanpa merasa malu saat melakukan kesalahan*" / I can speak alone with AI without feeling embarrassed when making mistakes). This low-anxiety environment encouraged more frequent practice, which in turn built confidence. Several respondents (e.g., NF, NA, LE, NL) noted they had begun to use English with their students and colleagues, demonstrating a transfer of confidence from the AI tool to real-world interactions.

The feature also most frequently cited as helpful was the detailed, immediate feedback on pronunciation, fluency, and vocabulary (e.g., Respondent PR: "*bisa*

melihat apakah pronunciationnya kita sudah benar atau tidak" / I can see if our pronunciation is correct or not; Respondent LE: "*fitur penilaian pelafalan, intonasi yang begitu detil*" / the feature for assessing pronunciation and intonation which is very detailed). This objective evaluation allowed users to identify specific areas for improvement, moving beyond subjective self-assessment. The quantitative scores (e.g., "score 6/9") provided in the logbooks served as tangible markers of progress, further motivating learners.

Data analysis also shows the relevance of content to the daily context. The relevance of the practice topics (e.g., Introduction and Daily Activities, Asking Permission and Giving Suggestions, and Describing Food,) to the respondents' daily lives in the pesantren was a critical success factor. Respondent SA explicitly stated, "*materi tes SpeechAce sesuai dengan agenda sehari-hari di pesantren*" (SpeechAce test material is in accordance with the daily agenda in the pesantren). This contextual relevance ensured that the acquired vocabulary and phrases were immediately applicable, enhancing both learning retention and practical utility.

Data analysis also reveal that AI SpeechAce facilitated self-directed learning beyond the classroom. Respondents appreciated the ability to practice autonomously at their own pace and time (e.g., Respondent AR: "*Bisa mulai belajar mandiri walau di luar kelas*" / I can start independent learning even outside the class; Respondent IN: "*bisa latihan di luar kelas juga*" / I can also practice outside the class). Such testimonies indicate that SpeechAce supports independent learning habits. This flexibility proved particularly beneficial for busy teachers balancing professional and personal demands.

Challenges in the Implementation of AI-Based Practice

Despite the overall positive outcomes, the data also highlights significant challenges that hinder optimal implementation. The most commonly reported challenges were related to core language proficiency, including the lack of vocabulary, which was cited by a majority of respondents such as M, SI, PR, MR, and NA as a main challenge that hindered fluency and the ability to formulate responses quickly. Pronunciation difficulties also posed a considerable challenge, with many respondents, including HR, NL, KA, and MZ, finding it challenging to articulate unfamiliar words or complex sentences; respondent MZ likened the experience to a tongue-twister ("*Kayak patah lidah*"). Additionally, grammar inconsistencies remained a persistent challenge for several participants, such as UF and PR, particularly concerning the correct use of verb forms (V1, V2, and V3).

Another challenge mentioned relates to the technical issue when using the SpeehAce. Multiple respondents (e.g., NSN, SI, NA, MR) explicitly suggested that the tool should be "unlimited" or that the "link should not be quick to limit." One important consideration is that this study employed the free version of SpeechAce rather than the premium package. As a result, full access to the platform was limited. To address this, participants were instructed to inform the researchers whenever a test link could no longer be accessed so that the researchers could send them a new link to get to the test platform. Other technical issues like loading times

and difficult access were also noted (e.g., NL, HR), which can disrupt the learning flow and discourage consistent use. Such limitations highlight the need for improved accessibility in future implementations.

Moreover, while the AI was effective for independent practice, applying these skills in spontaneous, real-life conversations with people remained a daunting challenge for many. Respondents NTS and W identified "*mempraktekkan bahasa Inggris langsung dengan orang lain*" (practicing English directly with others) and "*berbicara spontan*" (spontaneous speaking) as their main challenges. This indicates that while AI is an excellent tool for promoting confidence and skills, it may not fully replicate the unpredictability of human interaction. This underscores the gap between AI SpeechAce simulated practice and authentic communication. AI SpeechAce thus serves as a foundation, but additional real-life speaking opportunities remain essential.

Discussion: Implications and Future Directions

The findings of this study provide meaningful insights into the potential of AI SpeechAce in supporting teachers' speaking for daily communication skills, particularly in Islamic Boarding School context. Consistent improvements in vocabulary and grammar show that AI SpeechAce can help teachers enrich their vocabulary, apply more accurate sentence structures, and gain confidence in daily communication. The immediate, individualized feedback offered by SpeechAce created a low-anxiety learning environment (Abdelhalim & Alsehibany, 2025) that enable teachers to transfer their practice directly into real-life interactions with colleagues and students. This aligns with previous studies that highlight AI's potential in reducing anxiety (Miftahul et al., 2024), boosting self confidence to practice speaking English (Celik et al., 2025), promoting self-regulated learning (Mohebbi, 2025) which eventually enhance vocabulary and grammar aspects in speaking skills (Khalil et al., 2025; Mudhsh et al., 2025; Shaheen et al., 2024).

Besides, the topics learned and tested were closely related to daily communication in the boarding school context, making them directly relevant and meaningful for the teachers' real-life interactions. This finding is in line with Abdallah (2025), who emphasizes that understanding and responding to language learners' diverse needs. He argues that tailoring instruction to learners' contexts fosters more effective and meaningful communication. By introducing topics and situations that are familiar and directly relevant to the teachers' context, this approach increases both engagement and the practical transfer of learning.

The study also underscores the limits of relying solely on AI. Gains in fluency and pronunciation were modest and tended to be stagnant, with teachers reporting persistent difficulties in spontaneous speech and accurate pronunciation (Zou et al., 2024). These outcomes indicate that AI SpeechAce practice, though valuable, may not fully address the complexities of spoken interaction. A more structured AI based practice combined with conventional teaching instruction is desirable. This suggests that AI can complement, but not replace, teachers' roles (Rusdi et al., 2024; Üretmen Karaoğlu & Doğan, 2025; Yang, 2025).

Several limitations should be acknowledged. The study involved a small, homogenous group of teachers from a single pesantren which limit the generalizability of the findings. It also focused exclusively on one specific AI tool, SpeechAce, without comparing alternative platforms. Furthermore, the short duration of the intervention (only a month) restricted the ability to assess long-term impacts. These limitations point to clear directions for future research. Larger, more diverse samples across different school types would test the validity of the results. Longitudinal studies could confirm whether AI-assisted speaking tools are sustained over time. Comparative studies involving multiple AI platforms would also provide richer insights into effective pedagogical instructions.

4. Conclusion

In conclusion, the effectiveness of AI SpeechAce in enhancing speaking for daily communication skills of teachers at an Islamic Boarding School in Lhokseumawe, Aceh, is heavily influenced by its ability to create a safe, feedback-rich learning environment that boosts confidence and is relevant to their daily communicative needs. The primary challenges are not related to the AI itself, but rather a combination of teachers' pre-existing linguistic gaps and technical issues. For future research, a more structured blended learning model and a longer intervention are recommended. Since this study already combined AI practice with teaching sessions, future work could focus on designing more comprehensive integration between automated feedback and longer classroom intervention. This would better connect accuracy-building with fluency and spontaneity, ensuring that AI-supported learning transfers more effectively into real-world communication.

Acknowledgement

This research was supported by the National Competitive Grant Research Program, namely Penelitian Dosen Pemula (PDP), funded by the Directorate of Research, Technology, and Community Service, Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia (Kemendikbudristek RI). The author gratefully acknowledges this support, which made the research possible.

References

- Abdallah, M. M. (2025). Realistic needs matter: the imperative of learner-centred design in EFL curriculum development. *Online Submission*. <https://files.eric.ed.gov/fulltext/ED672712.pdf>
- Abdelhalim, S. M., & Alsehibany, R. A. (2025). Integrating AI-powered tools in EFL pronunciation instruction: effects on accuracy and L2 motivation. *Computer Assisted Language Learning*, 1-25. <https://doi.org/10.1080/09588221.2025.2534015>
- Alam, M. Z., Rahman, M. M., & Mizan, F. B. (2024). Teaching L2 speaking: A case study of English for academic purpose teachers' cognition and practices. *Mextesol Journal*, 48(3), n3. <https://eric.ed.gov/?id=EJ1446577>
-

-
- Amiruddin, M., & Jannah, U. R. (2019). Peran lingkungan bahasa dalam pemerolehan bahasa Inggris lisan santri di Pondok Pesantren Nurul Jadid Paiton. *Jurnal Pendidikan Edutama*, 6(1), 65-76. <http://ejurnal.ikipgribojonegoro.ac.id/index.php/JPE>
- Celik, B., Yildiz, Y., & Kara, S. (2025). Using ChatGPT as a virtual speaking tutor to boost EFL learners' speaking self-efficacy. *Australian Journal of Applied Linguistics*, 8(1), 102418. <https://doi.org/10.29140/ajal.v8n1.102418>
- Creswell, J. W. (1999). *Mixed-method research: Introduction and application*. In Handbook of educational policy (pp. 455-472). Academic press.
- Dong, W., Pan, D., & Kim, S. (2024). Exploring the integration of IoT and Generative AI in English language education: Smart tools for personalized learning experiences. *Journal of Computational Science*, 82, 102397. <https://doi.org/10.1016/j.jocs.2024.102397>
- El Shazly, R. (2021). Effects of artificial intelligence on English speaking anxiety and speaking performance: A case study. *Expert Systems*, 38(3), e12667. <https://doi.org/10.1111/exsy.12667>
- Karim, S. A., Hamzah, A. Q. S., Anjani, N. M., Prianti, J., & Sihole, I. G. (2023). Promoting EFL students' speaking performance through ELSA Speak: An artificial intelligence in English language learning. *JOLLT: Journal of Languages and Language Teaching*, 11(4), 655-668. <https://doi.org/https://doi.org/10.33394/jollt.v11i4.8958>
- Khalil, M., Amreen, S., & Ather, F. (2025). Exploring the role of AI chatbots in enhancing English grammar and vocabulary acquisition: A study of learning dynamics, user engagement, and linguistic development. *Journal of Applied Linguistics and TESOL (JALT)*, 8(1), 801-812. <https://jalt.com.pk/index.php/jalt/article/view/388/313>
- Mardhiah, A., Purwati, D., & Muna, H. (2024). Artificial Intelligence (AI) in enhancing spoken English proficiency: A systematic literature review. *International Journal of English Linguistics, Literature, and Education (IJELLE)*, 6(2). <https://doi.org/10.32585/ijelle.v6i2.5980>
- Masrukhan, L., Sa'diyah, M., Bahrudin, E., Dewi, C., & Bendriyanti, R. P. (2023). Evaluasi kinerja guru bahasa inggris guna meningkatkan mutu pendidikan di madrasah aliyah. *Paedagogia: Jurnal Kajian, Penelitian dan Pengembangan Kependidikan*, 14(3), 379-388.
- Mohebbi, A. (2025). Enabling learner independence and self-regulation in language education using AI tools: a systematic review. *Cogent Education*, 12(1), 2433814. <https://doi.org/10.1080/2331186X.2024.2433814>
- Miftahul, F. S., Zahria, A., & Rizka, M. A. (2024). Enhancing language learners' speaking skills and confidence with AI: A study of adaptive interactive methods. In *International Conference on Education, Science, Technology and Health (ICONESTH)* (pp. 1068-1079). <https://eproceeding.bbg.ac.id/iconesth/article/view/513>
- Mudhsh, B. A., Muqaibal, M. H., Al-Maashani, S., & Al-Raimi, M. (2025). Utilization of artificial intelligence tools in fostering English grammar and vocabulary among Omani EFL learners. *World*, 15(5). <https://doi.org/10.5430/wjel.v15n5p51>
-

- Ningsih, F. (2024). Analyzing students' English-speaking skills using speechace: Insights from an AI-powered assessment tool. *Pedagogic Research-Applied Literacy Journal*, 1(3), 111-124. <https://doi.org/10.70574/9w2prx09>
- Rahman, F., Al-Nahdi, A., & Zurriyati, Z. (2023). The (in)famous English language policy in Pesantren: What we already know and what remains unknown. *Idarah: Jurnal Pendidikan dan Kependidikan*, 7(2), 165–182. <https://doi.org/10.47766/idarrah.v7i2.821>
- Rusdi, L., Mardhiah, A., Muna, H., & Fitriani, S. (2024). Teacher's role in facilitating EFL autonomous learning through digital assistance: A systematic review. *Proceeding of ICELT International Conference on English Language Teaching*. 10-26. <https://journal.ar-raniry.ac.id/ICELT>
- Shaheen, F., Hussain, K. G., Jahan, M., & Azam, L. A. (2024). Exploring artificial intelligence's role in second language vocabulary acquisition: Insights from students' and teachers' perspectives. *Al-Qanṭara*, 400-419. <https://alqantarajournal.com/index.php/Journal/article/view/580>
- Üretmen Karaoğlu, S., & Doğan, C. (2025). EFL teachers' insights on incorporating AI in language education. *Journal of Theoretical Educational Science/ Kuramsal Eğitim Bilim Dergisi*, 18(3).
- Yang, M. L. (2025). Artificial intelligence versus human teacher assistants and language learners' progress in learning and retention of complex sentences. *Current Psychology*, 44(7), 6292-6304. <https://doi.org/10.1007/s12144-025-07488-6>
- Yeh, H. C. (2025). The synergy of generative AI and inquiry-based learning: transforming the landscape of English teaching and learning. *Interactive Learning Environments*, 33(1), 88–102. <https://doi.org/10.1080/10494820.2024.2335491>
- Zhang, X., Sun, J., & Deng, Y. (2023). Design and application of intelligent classroom for English language and literature based on artificial intelligence technology. *Applied Artificial Intelligence*, 37(1), 2216051. <https://doi.org/10.1080/08839514.2023.2216051>
- Zou, B., Guan, X., Shao, Y., & Chen, P. (2023). Supporting speaking practice by social network-based interaction in artificial intelligence (AI)-assisted language learning. *Sustainability*, 15(4), 2872. <https://doi.org/10.3390/su15042872>
- Zou, B., Liviero, S., Ma, Q., Zhang, W., Du, Y., & Xing, P. (2024). Exploring EFL learners' perceived promise and limitations of using an artificial intelligence speech evaluation system for speaking practice. *System*, 126, 103497. <https://doi.org/10.1016/j.system.2024.103497>

How to cite this article:

Mardhiah, A., Hajar, I., Purwati, D., Muna, H., & Ariani, R. (2025). Leveraging Artificial Intelligence (AI) to Enhance Teachers' Spoken English at an Islamic Boarding School. *Journal of Educational Sciences*, 9(6), 5163-5176.